

C  
C  
M  
R  
P

THE CALIFORNIA REPORT ON  
CORONARY ARTERY  
BYPASS GRAFT SURGERY

1999 HOSPITAL DATA

TECHNICAL APPENDIX: AUDIT SUMMARY 2003



California CABG Mortality Reporting Program  
March 2004

This report is a product of the California Coronary Artery Graft (CABG) Mortality Reporting Program (CCMRP), a project sponsored by the Pacific Business Group on Health (PBGH) and the California Office of Statewide Health Planning and Development (OSHPD). PBGH is a non-profit organization of large employers in California who work to provide independent information to consumers on the quality of care provided by health plans, hospitals, and doctors (see [www.healthscope.org](http://www.healthscope.org)).

OSHPD is the state department that produces risk-adjusted hospital outcomes data  
(see [www.oshpd.state.ca.us](http://www.oshpd.state.ca.us)).

Pacific Business Group on Health and Office of Statewide Health Planning and Development,  
March 2004

# THE CALIFORNIA REPORT ON CORONARY ARTERY BYPASS GRAFT SURGERY

1999 Hospital Data

## Technical Appendix: Audit Summary

**March 2004**

This report was prepared by:

Beate Danielsen, Ph.D.  
Joseph P. Parker, Ph.D.  
Cheryl L. Damberg, Ph.D.  
Anne Castles, M.A., M.P.H.  
Anthony E. Steimle, M.D.

Suggested citation: Danielsen, B, Parker, JP, Damberg, CL, Castles, AG, and Steimle, AE. ***The California Report on Coronary Artery Bypass Graft Surgery 1999 Hospital Data: Technical Appendix - Audit Summary***, San Francisco, CA: Pacific Business Group on Health and the California Office of Statewide Health Planning and Development, March 2004.

Additional copies of the Technical Report can be obtained through the PBGH ([www.pbgh.org](http://www.pbgh.org)) and OSHPD ([www.oshpd.state.ca.us](http://www.oshpd.state.ca.us)) Web sites. PBGH posts the hospital performance results on its California Consumer HealthScope Web site ([www.healthscope.org](http://www.healthscope.org)), a public source of information on healthcare quality for California consumers.



## TABLE OF CONTENTS

INTRODUCTION .....	5
DESIGN OF THE 1999 AUDIT .....	5
Selection of Hospitals for Audit .....	5
Record Selection within Hospitals .....	6
Variables for Audit .....	7
Reserve Records.....	7
AUDIT PROCESS .....	7
Hospital Participation .....	7
Identification of Non-isolated CABGs .....	8
Unexpected Problems .....	8
Rater Reliability.....	8
COMPARISON OF 1999 AND 1997/8 AUDIT .....	9
ANALYSIS OF AUDITED DATA.....	9
Bivariate Frequency Analysis .....	9
Agreement Analysis .....	10
Sensitivity Analysis .....	18
KEY FINDINGS FROM THE AUDIT OF HOSPITAL DATA.....	22

## TABLES AND FIGURES

### TABLES

Table 1: Comparison of Audited Data and CCMRP Submissions for Acuity, All Hospitals, 1999 Data .....	10
Table 2: Agreement Statistics, All Hospitals, 1999 Data.....	11
Table 3: Comparison of Audited Data and CCMRP Submissions for Discharge Status, All Hospitals, 1999 Data.....	13
Table 4: Comparison of Audited Data and CCMRP Submissions for NYHA CHF Class, All Hospitals, 1999 Data.....	13
Table 5: Comparison of Audited Data and CCMRP Submissions for CCS Angina Class, All Hospitals, 1999 Data.....	13
Table 6: Comparison of Audited Data and CCMRP Submissions for Myocardial Infarction, All Hospitals, 1999 Data.....	14
Table 7: Agreement Measures for Example Hospital, 1999 Data .....	18
Table 8: Impact on Fitted Model Characteristics when Replacing Audited Records with Information from Audit, 1999 Data .....	19
Table 9: Rating Changes after Replacement with Audited Data, 1999 Data .....	21

### FIGURES

Figure 1: Creatinine Agreement Chart, All Hospitals, 1999 Data.....	15
Figure 2: Acuity Agreement Chart, All Hospitals, 1999 Data .....	16
Figure 3: NYHA CHF Class Agreement Chart, All Hospitals, 1999 Data.....	17

## INTRODUCTION

The audit report complements the 1999 hospital and technical report. This document describes the audit that was performed to verify the integrity of hospital reported data, confirm hospital outlier status, identify coding problems, and study the implications of the missing value assignment strategy adopted for the 1999 data analysis.

After a description of the audit design including hospital selection, record selection within each hospital, and the list of variables audited, the audit process and subsequent data analysis are summarized. A list of key findings from the audit of hospital data concludes this report.

For all variables reviewed in the audit, with the exception of Discharge Status (alive/dead), the results were considered the “gold” standard and were used in place of data submitted by the hospital in conducting the final analysis of hospital performance results. For Discharge Status, hospitals were contacted to make corrections when discrepancies were found between what the hospital submitted and what was found either through the audit or the link to the OSHPD Patient Discharge Data (PDD). Given the importance of the key outcome variable in determining a hospital’s performance ranking, hospitals were asked to provide substantiating evidence when discharge status differed from the audit or PDD result.<sup>1</sup>

At the conclusion of the audit, all audited hospitals were sent detailed reports showing how well they coded and were given an opportunity to make additional data corrections.

## DESIGN OF THE 1999 AUDIT

Thirty-one of the 70 hospitals that chose to publicly report results for 1999 underwent an audit. Note that a total of 36 hospitals were audited, however five of those withdrew from participation after the audit. Selection of the audit sample entailed two steps: 1) determining which hospitals would be audited, and 2) determining which records would be audited in each of the selected hospitals.

### Selection of Hospitals for Audit

After the data cleaning and initial analyses were completed, two preliminary statistical models for the mortality outcome were developed. Model 1 was developed using the **1999 Data** submission only; Model 2 was developed using the **All Quarters** data (1997-1999). Based on the two models’ predictions, we identified above (“better than expected”) and below average (“worse than expected”) performers. A hospital was considered an outlier if it reported fewer or more deaths than would be expected based on the hospital’s patient case-mix. The probability at which we considered a hospital an outlier was 0.126 (two-tailed) for the **1999** model (Model 1) and 0.01 (two-tailed) for the **All Quarters** model (Model 2). The probability level for 1999 was intentionally set higher to ensure that hospitals found to be outliers at a statistically significant level (0.05 or less) and those hospitals just above the statistically significant level (“close to the

---

<sup>1</sup> While a subset of hospitals were subject to the audit, 100% of participating hospitals had their data linked to the OSHPD PDD to confirm the coding of discharge status and whether the isolated case submission was complete or included potentially non-isolated cases.

fence”) were audited. In addition to selecting hospitals that performed “worse” or “better” than expected, the 1999 data audit also included a random sample of hospitals that were classified as “no different than expected,” with the following exceptions:

- Hospitals that went out of business after 1999.<sup>2</sup>
- Hospitals that withdrew from CCMRP prior to the audit.<sup>3</sup>
- Hospitals that were not identified as outliers for 1999 and that were audited for the 97/98 data submission.

After the preliminary analyses, 18 hospitals were selected for audit based on their outlier status (11 “worse than expected” and 7 “better than expected”). An additional 20 hospitals were selected at random from the remaining hospitals subject to the exceptions listed above. Thirty-eight hospitals representing 52% of CCMRP submissions were initially targeted for audit, and 23% of all records submitted by these hospitals were requested for audit.

Two hospitals that had originally submitted data for analysis and were selected for audit refused to be audited. Because the audit is a condition of participation, these hospitals were dropped from the program because we could not confirm the accuracy of their data submission.

### **Record Selection within Hospitals**

Once the hospitals were selected for audit, the audit sample for each hospital was designed based on the following considerations.

**Overall Sample Size:** The total number of records to be sampled was limited by budgetary constraints, which required balancing the total number of hospitals and records that could be reviewed. The total number of records at all audited hospitals was not to exceed 3,200 records (31% of records submitted).

**Sample Type:** To make statistically meaningful statements about the differences between hospital-submitted and audited data, it was desirable to draw a random sample of records. At the same time, CCMRP wanted to ensure that records pertaining to more seriously-ill patients with more complex records were adequately selected. A simple random sample would not likely include sufficient representation of the severely-ill. CCMRP elected to draw a weighted random sample. First, all deaths were selected for audit. Among the remaining records, the predicted risk of death for each patient was used as a sampling weight. Using this approach, records pertaining to more seriously ill patients were more likely to be audited than records pertaining to less seriously ill patients. The effect of this biased sampling scheme is observed in comparing the mean expected mortality for all records submitted by audited hospitals (2.6%) versus the mean expected mortality for the records requested for audit (5.7%).

**Sample Size for Each Audited Hospital:** Audit samples were pulled proportional to each hospital’s volume of isolated CABG surgeries, adding a constraint that each hospital sample contain a minimum and maximum number of records. Pulling a sample proportional to hospital volume allows for more accurate statistical statements about differences between records submitted by the hospital and audited records. Because a minimum number of cases is needed to produce statistically meaningful results for small-volume hospitals, the minimum number of

---

<sup>2</sup> The Heart Hospital of the Desert and Long Beach Community Hospital closed their CABG programs after 1999.

<sup>3</sup> These hospitals include Huntington Memorial Hospital, Good Samaritan Hospital-San Jose, and St. Agnes Medical Center.

cases per hospital was 40. Setting a maximum alleviated any unnecessary burden on large-volume hospitals, which would otherwise be required to pull a very large number of cases. The maximum number of cases was 160 per audited hospital.

### Variables for Audit

Recognizing the importance of some variables compared to others with respect to the mortality outcome, the following variables were selected for the 1999 data audit.

Isolated CABG (Yes/No)	Chronic Obstructive Pulmonary Disease (COPD)	Hypertension	Dialysis
Peripheral Vascular Disease (PVD)	Cerebrovascular Disease (CVD)	Ventricular Arrhythmia	Myocardial Infarction
Time since Most Recent Myocardial Infarction	Creatinine prior to Surgery	Number of Prior Cardiac Operations	Congestive Heart Failure (CHF)
Acuity	Diabetes	NYHA CHF Angina Class	Angina
Angina Type	CCS Angina Class	Ejection Fraction	Ejection Fraction Measurement Method
Left Main Stenosis	PTCA on Current Admission	Interval between PTCA and Surgery	Discharge Status
Date of Death (if died)	Date of Discharge		

### Reserve Records

A second set of “reserve” records within the set of records requested for audit from each hospital were designated for review in the event that requested Priority 1 records could not be used for any reason (e.g., unable to locate record, the auditors determined that the record was not an isolated CABG). Reserve or “priority 2” records were also sampled using the probability of death as sampling weight. For hospitals with 100 or fewer CCMRP submissions, 10 extra records were pulled; for hospitals with more than 100 CCMRP submissions, 20 extra records were pulled.

## AUDIT PROCESS

CCMRP contracted with the Health Services Advisory Group (HSAG) to conduct the independent, external audit. HSAG is an Arizona-based peer-review organization with experience in abstracting cardiovascular information from medical records. The audit was carried out between March and September 2001. Six RN abstractors from HSAG attended a training class using the same training materials employed at the hospital training sessions.

### Hospital Participation

Two hospitals, Mercy San Juan Hospital and Mercy General Hospital, chose not to participate in the audit. These two hospitals accounted for 206 audit records requested, including eight deaths. Deleting the two hospitals that declined to be audited, the hospitals selected for audit submitted a total of 10,273 records to CCMRP, representing approximately 46% of all records submitted. CCMRP selected 2,472 or 24% of these records for audit, including 306 deaths. Overall, the auditors were able to review 97.4% of records requested. The percent of records

audited at individual hospitals ranged from a minimum of 88.7% to a maximum of 109% of the records requested.

### **Identification of Non-isolated CABGs**

A total of 2,408 cases were audited. Forty-four records (1.8%) were determined to be non-isolated CABGs by the auditors. A second review of these cases by CCMRP's medical advisor resulted in the removal of all but five of the cases from the analytic file. CCMRP also used OSHPD's Patient Discharge Data to verify the information collected by the auditors.

### **Unexpected Problems**

A complication arose because auditors did not understand the distinction between priority 1 and priority 2 records. Therefore, for some hospitals a large number of priority 1 records – including records pertaining to in-hospital deaths – were not audited. Overall, 84.8% of priority 1 and 79.5% of death records requested for audit were audited. The percent of priority 1 records audited ranged from 76% to 100%. Consequently, for some hospitals the average severity of audited records was lower than the average severity of the records requested for audit.

### **Rater Reliability**

The audit vendor used physician over-reads on a subset of cases CCMRP selected for audit. Dr. Carol Ann Zaher, Chief of Cardiology at Kaiser Permanente Medical Center from 1983-1996, was hired to perform a re-review (over-read) of 2% of the records abstracted. Dr. Zaher performed an over-read of 44 records during an on-site visit at two of the hospitals selected for audit. Dr. Zaher's over-read occurred at the onset of the audit in order to clarify any misunderstandings of the tool or misinterpretation of the data at the earliest possible stage. Dr. Zaher re-reviewed records across all auditors and provided feedback and training immediately.

In addition to the physician over-read at the onset of the audit, 5% percent of the abstracted records at each audited facility were randomly selected for auditor re-review in order to validate the accuracy of the data collection. The review process consisted of one auditor (Auditor #1) initially abstracting the record. A second auditor then randomly pulled 5% of the records abstracted by the first auditor and performed a second independent audit. Auditor #2 then compared her responses to the ones abstracted by Auditor #1. When discrepancies arose, the two auditors would discuss the issue and determine the appropriate coding by referring back to the medical record. If unable to agree, the auditors contacted the project leader for a final decision. Once the audited records were received at HSAG, they were reviewed for accuracy and completeness. Auditors were contacted by phone to resolve any remaining discrepancies prior to processing.

Using Cohen's Kappa to assess inter-rater reliability, the overall reliability for the audit was 97.58% and ranged from 95% to 99% for the individual auditors.<sup>4</sup> A total of 165 or 6.8% of the records abstracted were over-read.

---

<sup>4</sup> Cohen, JA. A coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20 (1), 37-46.

## COMPARISON OF 1999 AND 1997/8 AUDIT

Using the 1997/1998 CCMRP data audit as a learning experience, the design of the 1999 data audit was improved in several regards.

For the 1997/98 data collection period, 33% of the 79 CCMRP participants were audited, whereas for the 1999 data audit, 36 of the 84 hospitals (43%) eligible for audit were audited.

In contrast to the 1997/1998 data audit which was restricted to “outlier” hospitals, the 1999 data audit included not only all outlier hospitals, but also institutions that were randomly selected from the group of non-outlier hospitals (i.e., those performing no differently than expected).

For the 1997/1998 data audit, 40 records were pulled from each audited hospital. For 1999, CCMRP adopted a proportional sampling strategy. Approximately 24% of records submitted by a hospital were selected for audit, with a minimum of 40 and a maximum of 160 records from each hospital.

For the 1997/1998 data audit, all deaths and high severity cases were systematically pulled. For the 1999 data audit, all deaths were systematically pulled as well; however the remaining records were pulled based on a weighted random sample. The weight used in the sample was based on the predicted probability of death of each case based on a preliminary in-hospital mortality model.

Because CCMRP increased both the number of hospitals and records audited for 1999, audit costs increased from \$132,000 to \$350,000.

## ANALYSIS OF AUDITED DATA

The audited data were compared to the data that hospitals originally submitted to CCMRP. It is important to understand that the CCMRP data on which the audit sample was based differed from the final CCMRP data that was used for the generation of the 1999 Hospital Report on CABG surgeries. After the audit sample was determined, hospitals continued to submit, change, and withdraw records. In order to compare the results of the audit to the most appropriate representation of CCMRP data, all CCMRP data used in this report are those as they were available in June 2001.

### Bivariate Frequency Analysis

Bivariate frequencies were generated and analyzed in order to evaluate the accuracy of hospital coding for each variable.

For each audited variable, we produced a bivariate table that compared audit-obtained data to that in the original CCMRP submission. Table 1 displays, for example, the bivariate frequencies of *Acuity*. The table represents data for all audited hospitals combined. Records with missing information are not included in the table. Rows refer to the CCMRP coding; columns refer to the coding according to the audit. For all shaded cells that lie on the diagonal CCMRP submission and audit agree, whereas for those not shaded and not on the diagonal, CCMRP submission and audit disagree.

To understand the implications of the audit, consider the cells representing actual categories (i.e., elective, urgent, emergent, salvage) of the variable to be a rectangular box. Cells that lie in the upper right triangle of the box represent cases where down-coding occurs (i.e., the CCMRP coding incorrectly characterizes the patient as lower risk and will produce a less favorable hospital score). Table 1 reveals that the audit classified 439 cases as urgent, emergent, or salvage where hospitals had originally submitted elective. Cells in the lower left triangle of the box represent cases where up-coding occurs (i.e., the coding incorrectly characterizes the patient as higher risk, which may produce a more favorable hospital score).

In Table 1, the diagonal is highlighted. For these cells, CCMRP and audited data agreed. Overall, almost 66% of records audited agreed on acuity.

**Table 1: Comparison of Audited Data and CCMRP Submissions for Acuity, All Hospitals, 1999 Data<sup>5</sup>**

		Audited Data				Total
		Elective	Urgent	Emergent	Salvage	
CCMRP Data	Elective	447	431	7	1	886
	Urgent	140	911	53	4	1,108
	Emergent	16	117	199	3	335
	Salvage	1	18	29	4	52
Total		604	1,477	288	12	2,381

## Agreement Analysis

In addition to generating these bivariate tables for each audited variable for each hospital, CCMRP also tabulated the following statistics (see Table 2):

- Percent of missing values that would have been incorrectly assigned to the default low risk category in the absence of audited data.<sup>6</sup> Tabulation of the percent of missing values that were incorrectly assigned to the low risk category allows an assessment of how much under-coding this policy implies for hospitals that have incomplete data.
- Percent of records in agreement: The percent of records in agreement can be used as a first indicator of reliability of coding. However, note that for variables that describe rare events (such as death), the percent of records in agreement can be high, however, even small levels of disagreement can imply large changes in a hospital's rating.
- Severity weighted percent of records in disagreement based on the lower left triangle of the bivariate frequencies: As noted above, the cells in the lower left triangle represent records that appear to be up-coded. This statistic thus reflects the amount and severity of up-coding as a percentage of the overall amount and severity of coding disagreement for the variable. The statistic is based on both the number of disagreements and the severity of the disagreement. Specifically, CCMRP weighted the number of disagreements proportionally by the severity of the coded category in comparison to the reference (low risk) category. Weighting is used to reflect the fact that some coding inaccuracies are

<sup>5</sup> The numbers in the table reflects the data used at the time the final audit analysis was performed in September 2001. However data corrections made by hospitals after September 2001 may slightly alter these numbers.

<sup>6</sup> CCMRP's policy is to substitute missing variables with the lowest risk category for the variable.

more egregious than others (elective cases coded as salvage will lead to a more biased result than elective cases coded as urgent). Note that choosing equidistant or severity related weights does not matter for binomial outcomes. However, it does matter for outcomes with more than two categories. Variables for which we do not observe any systematic up-coding or down-coding will have statistics in the range of 50%. Variables that are systematically up-coded will have statistics greater than 50%. If the severity weighted lower triangle disagreement exceeds 80% and at the same time the percent of records in agreement is below 70%, the up-coding might be substantial.

- Cohen's kappa: We calculated Cohen's kappa solely to present a measure that is commonly used in the inter-rater reliability literature. However, the use of Cohen's kappa is problematic, particularly if one or more marginal cell counts are zero. As the latter was frequently the situation for hospital-level tabulations, the value of this statistic is diminished (statistic not shown in the table).

Table 2 displays agreement measures for all hospitals combined for each audited variable. The second column shows the number of records for which the variable was audited.<sup>7</sup> The third column indicates how often a variable was missing in the CCMRP submission, and the fourth column indicates how often these missing values would have been replaced incorrectly by the lowest risk category. Column 5 shows the percent of records in agreement, and column 6 shows the severity-weighted disagreement in the lower triangle of the bivariate frequency table.

**Table 2: Agreement Statistics, All Hospitals, 1999 Data**

Variable	Records Audited	Missing Values	% Missing Values that Would be Incorrectly Classified	% Agreement	% Lower Triangle Severity Weighted Disagreement
Acuity	2,408	2	100.00	65.56	64.36
Angina Type (Stable/Unstable)	2,408	0	NA	65.37	34.73
Angina (Yes/No)	2,408	0	NA	86.21	42.47
CCS Angina Class	2,408	105	79.05	45.76	53.19
Congestive Heart Failure	2,408	31	38.71	82.23	32.94
COPD	2,408	6	0.00	86.34	73.25
Creatinine (mg/dl)	2,408	556	3.96	93.31	56.37
Cerebrovascular Disease	2,408	3	0.00	87.67	45.79
Dialysis	2,408	91	0.00	98.13	86.67
Diabetes	2,408	3	0.00	94.73	45.67
Ejection Fraction (%)	2,408	228	15.79	78.95	60.27
Method of measuring ejection fraction	2,408	406	0.00	74.34	Not Calculated
Hypertension	2,408	7	85.71	84.39	40.43
Time from PTCA to surgery	125	45	42.22	78.40	12.50
Left Main Stenosis	2,408	388	7.22	85.96	51.46
Myocardial Infarction	2,408	5	60.00	68.31	57.68
Myocardial Infarction (Yes/No)	2,408	134	38.06	82.97	55.12

<sup>7</sup> The time from PTCA to surgery was only audited if a PTCA preceded surgery.

Variable	Records Audited	Missing Values	% Missing Values that Would be Incorrectly Classified	% Agreement	% Lower Triangle Severity Weighted Disagreement
NYHA CHF Class	2,408	100	24.00	48.84	77.96
Operative Incidence	2,408	249	1.61	96.89	29.60
PTCA on this Admission	2,408	839	3.46	91.36	67.79
Peripheral Vascular Disease	2,408	5	40.00	87.00	62.94
Gender	2,408	2	0.00	100.00	0.00
Status of patient at discharge	2,408	1	0.00	97.59	Not Calculated
Ventricular Arrhythmia	2,408	107	1.87	94.68	80.47

The following conclusions can be drawn from the data presented above:

The policy of replacing the lowest risk category for missing values was problematic for *CCS Angina Class*, *MI (yes/no)*, and *NYHA CHF Class*. For these three variables, the number of missing values was relatively high (100 or more values missing) and the percent of incorrect replacements was greater than 20%.

Agreement was relatively poor (below 70%) for the following variables: *Acuity*, *Angina Type*, *CCS Angina Class*, *Myocardial Infarction*, and *NYHA CHF Class*.

For the variables with poor levels of overall agreement, severity-weighted disagreement in the lower left triangle of the frequency table was relatively high (above 60%) for *Acuity* and *NYHA CHF class*. In other words, these two variables tended to be up-coded. Conversely, severity-weighted disagreement in the lower left triangle of the frequency table was low for *Angina Type*.

As the outcome variable could not be studied by all of the methods above, we looked at the bivariate frequency of discharge status (Table 3). The behavior of the *NYHA CHF Class*, *Acuity*, *CCS Angina Class*, and *Myocardial Infarction* variables warranted a closer inspection of the actual bivariate frequencies (Table 4, Table 5, and Table 6).

**Table 3: Comparison of Audited Data and CCMRP Submissions for Discharge Status, All Hospitals, 1999 Data<sup>8</sup>**

		Audited Data			Total
		Missing	Dead	Alive	
CCMRP Data	Missing	0	0	1	1
	Dead	2	219	15	236
	Alive	39	1	2131	2,171
Total		41	2,147	220	2,408

**Table 4: Comparison of Audited Data and CCMRP Submissions for NYHA CHF Class, All Hospitals, 1999 Data**

		Audited Data					Total
		Missing	I	II	III	IV	
CCMRP Data	Missing	1	75	5	3	16	100
	I	12	803	31	22	58	926
	II	6	190	20	12	31	259
	III	8	324	37	33	80	482
	IV	9	312	33	55	232	641
Total		36	1,704	126	125	417	2,408

**Table 5: Comparison of Audited Data and CCMRP Submissions for CCS Angina Class, All Hospitals, 1999 Data**

		Audited Data					Total
		Missing	I	II	III	IV	
CCMRP Data	Missing	3	21	16	19	46	105
	I	4	109	35	34	81	263
	II	1	56	56	67	75	255
	III	8	145	96	196	338	783
	IV	13	76	68	134	711	1,002
Total		29	407	271	450	1,251	2,408

<sup>8</sup> The numbers in the table reflects the data used at the time the final audit analysis was performed in September 2001. However data corrections made by hospitals after September 2001 may slightly alter these numbers.

**Table 6: Comparison of Audited Data and CCMRP Submissions for Myocardial Infarction, All Hospitals, 1999 Data**

		Audited Data							Total
		Missing	No MI	MI, but when unknown	21 days ago	7 to 20 days ago	1 to 6 days ago	within 1 day	
CCMRP Data	Missing	0	2	0	1	1	1	0	5
	No MI	0	765	1	107	18	53	12	956
	MI, but when unknown	0	10	1	25	3	12	0	51
	21 days ago	0	95	1	333	24	39	6	498
	7 to 20 days ago	0	18	0	25	79	37	3	162
	1 to 6 days ago	0	58	0	42	50	362	23	535
	within 1 day	0	31	3	9	0	55	103	201
	Total	0	979	6	542	175	559	147	2,408

The coding confusion about the *NYHA CHF Class* variable is remarkable. The majority of records coded as class I, II, or III by the auditors were coded as a different class in the CCMRP submission. The findings were similar for the *CCS Angina Class* variable. Due to the unreliable coding of both *NYHA Class* and *CCS Class*, these variables were excluded from the final **1999** and **All Quarters** risk models.<sup>9</sup>

The coding of *Acuity* was of particular concern as this variable has the greatest effect on the risk of mortality seen in the risk model (Table 1). While the difference in risk of adverse outcome is small between the elective and urgent categories, the difference in risk is substantial when comparing the Elective/Urgent categories to the Emergent/Salvage categories. Table 1 indicates that 152 records (6.4%) were coded as emergent or salvage while the audit found these records to pertain to patients with only elective or urgent acuity (the lower triangle area). Conversely, 65 records (2.7%) were coded as elective or urgent while the audit classified these records as emergent or salvage (the upper triangle area).<sup>10</sup>

<sup>9</sup> The Society of Thoracic Surgeons, for the 1999 data submissions to the STS cardiac database, did not distinguish separate uses for CCS Class and NYHA Class. As a consequence, a large number of hospitals used NYHA Class to code Angina functional severity as opposed to using this variable to code for Congestive Heart Failure.

<sup>10</sup> Table 1 highlights an important feature of the severity-weighted disagreement statistic. The total number of cases in the upper right triangle (representing down-coding) exceeds the number of cases in the lower left triangle (representing up-coding). However, the severity-adjusted disagreement statistic indicates a tendency to up-code. This is because in calculating the severity weight, the distance between categories is assumed proportional to the increase in risk of death pertaining to the two categories. For instance, going from Elective to Salvage implies 11 times the risk of death for a "Salvage" patient compared to an "Elective" patient. On the other hand, going from Elective to Urgent implies 1.23 times the risk of death for an "Urgent" patient compared to an "Elective" patient. As the vast majority of coding errors in the upper right triangle represent less egregious errors, the severity-weight adjusted disagreement statistic is less affected by these down-coding errors than it is by the relatively large number of patients coded as "Emergent" or "Salvage" when they were actually "Elective" and "Urgent."

The tabulation for *Myocardial Infarction* (Table 6) does not indicate a level of confusion in coding quite as severe as for *NYHA CHF Class* or *CHF Angina Class*. Tendencies to up-code were confounded by equally strong tendencies to down-code.

Agreement charts were used to study coding problems at the hospital level. Figure 1 displays an agreement chart for *Creatinine*. The horizontal axis and vertical axis both range from 0 to 100%. The horizontal axis represents the percent of agreement. The vertical axis represents the percent of severity-weighted disagreement in the lower left triangle of the frequency table. As indicated above, a low level of agreement coupled with a low/high percentage of severity-weighted disagreement indicates tendencies of down-/up-coding. Each circle in the chart represents a CCMRP hospital that participated in the audit. The area of the circle is proportional to the volume of cases audited. In addition, three circles representing the entire group of institutions rated better than expected, worse than expected or no different than expected, respectively, were also included in the chart. The agreement chart for Creatinine is an example of a variable that is coded acceptably well. The agreement exceeds 80% at all hospitals.

**Figure 1: Creatinine Agreement Chart, All Hospitals, 1999 Data**

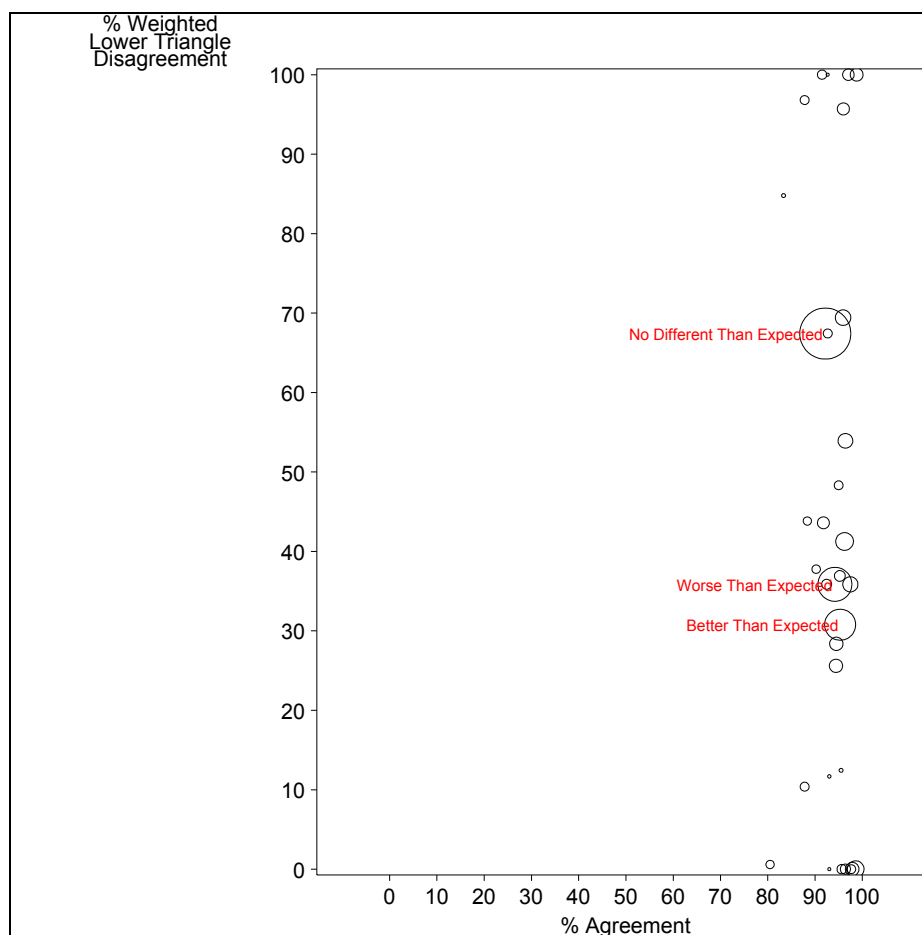
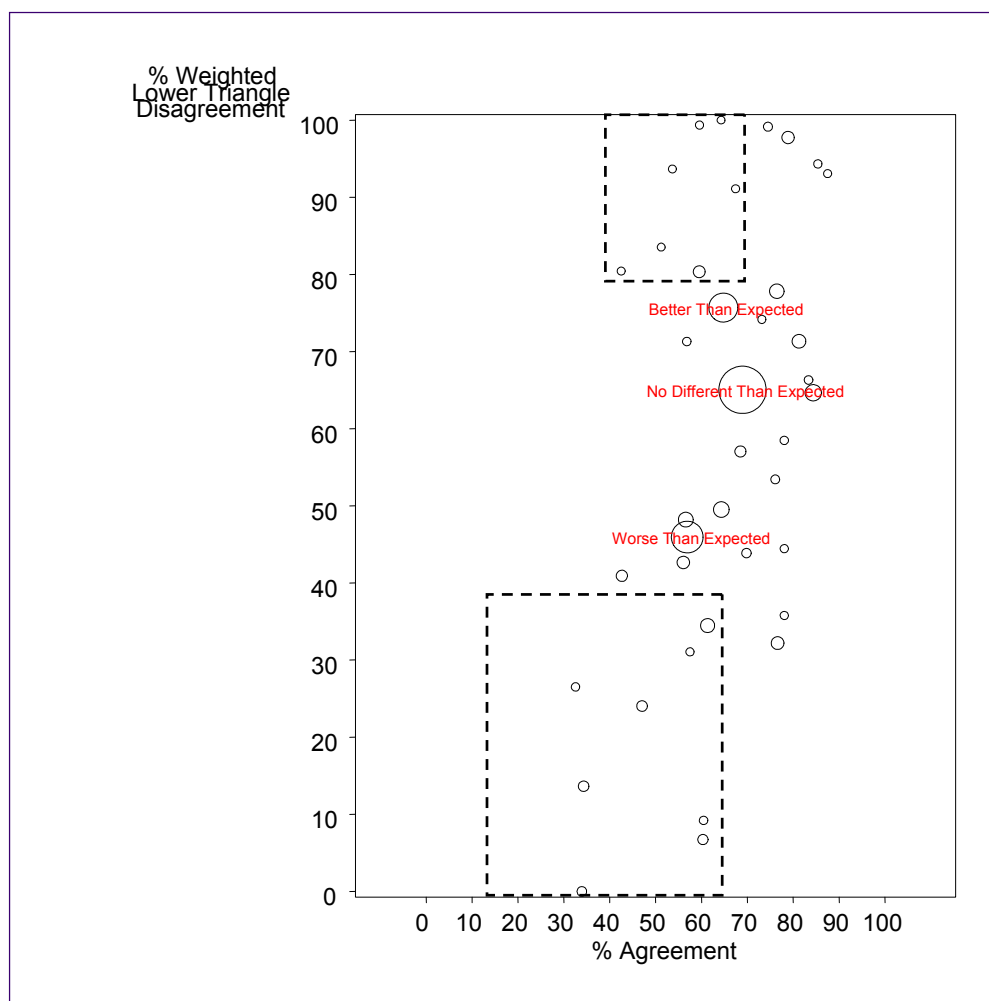


Figure 2 shows an agreement chart for *Acuity*. The dashed rectangles in the lower and upper portion of the chart identify institutions where coding was particularly problematic. For instance, for the upper rectangle, the level of agreement on Acuity between audit and CCMRP data is

under 70% while the severity-adjusted disagreement statistic exceeds 80%, indicating up-coding; for the lower rectangle, level of agreement on Acuity is under 70% while the severity-adjusted disagreement statistic is under 40% indicating down-coding.

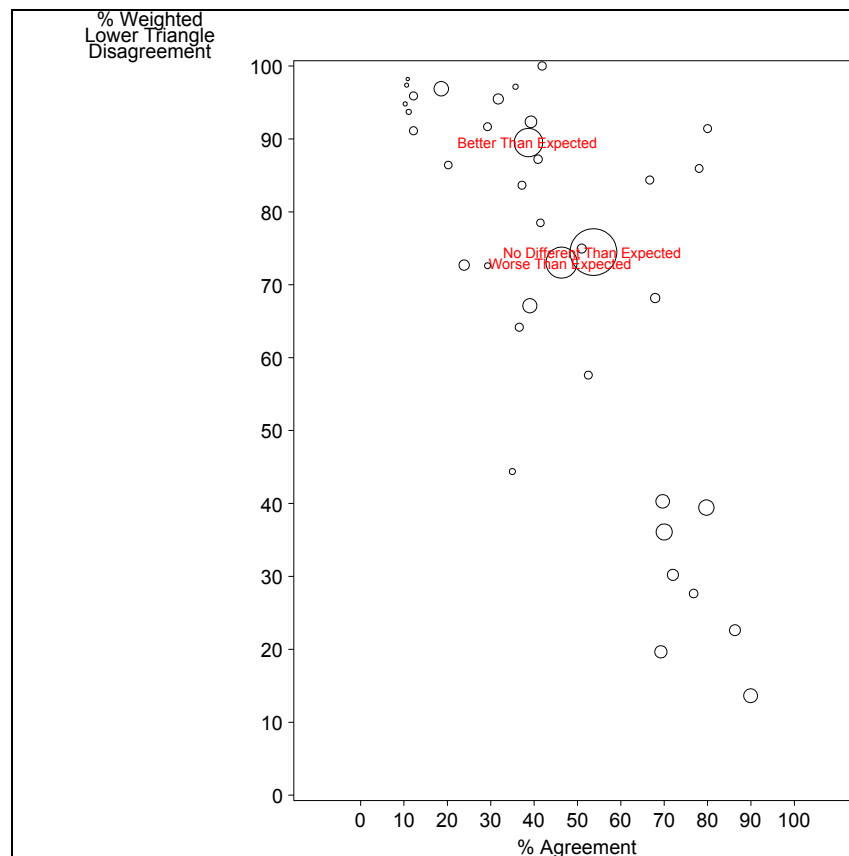
Hospitals that performed “worse than expected” tended to have lower levels of agreement on Acuity and lower levels of severity-adjusted disagreement. Hospitals that performed “better than expected” had a slightly higher level of agreement on Acuity and higher levels of severity-adjusted disagreement. These findings suggested that the classification of some of these hospitals as outliers may be a result of coding deficiencies. As such, hospitals identified in the analysis as under-coding acuity were given an opportunity to resubmit cases not audited with adjustments to the acuity variable. Hospitals engaged in clear up-coding of acuity were asked to re-submit their records with appropriate adjustments.

**Figure 2: Acuity Agreement Chart, All Hospitals, 1999 Data**



The agreement chart for *NYHA CHF Class* is shown in Figure 3. The chart indicates the presence of substantial coding problems of this variable at nearly all hospitals. An agreement chart for *CCS Angina Class* (not shown) led to the same conclusion.

**Figure 3: NYHA CHF Class Agreement Chart, All Hospitals, 1999 Data**



The audit results were also used to identify individual hospitals with coding problems. Table 7 shows an example of an individual hospital.<sup>11</sup> For this hospital, agreement on *Acuity*, *CCS Angina Class*, and *NYHA CHF Class* was poor. Agreement on *Myocardial Infarction* was acceptable. The severity-adjusted disagreement statistic suggests up-coding of the *Acuity* and *NYHA CHF Class* variables. In addition, for the *Myocardial Infarction* variable, all patient records in disagreement had been coded to a category indicative of higher severity than the audit found. For this institution, replacing the original hospital data with audited data resulted in a re-classification of its performance rating. The rating change occurred as a result of the miscoding of *Acuity*, as *CCS* and *NYHA Class* variables were omitted from the final risk model.

<sup>11</sup> To prevent identification of the hospital, we present percentages only.

**Table 7: Agreement Measures for Example Hospital, 1999 Data**

Variable	% Missing Values that Would be Incorrectly Classified	% Agreement	% Lower Triangle Severity Weighted Disagreement
Acuity	NA	51.22	83.54
Angina	NA	85.37	99.80
Angina (Yes/No)	NA	95.12	0.00
CCS Angina Class	57.14	53.66	40.59
Congestive Heart Failure	NA	82.93	28.57
COPD	NA	95.12	50.00
Creatinine (mg/dl)	NA	87.81	96.80
Cerebrovascular Disease	NA	85.37	66.67
Dialysis	0.00	100.00	.
Diabetes	NA	97.56	100.00
Ejection Fraction (%)	20.00	90.24	0.00
Method of measuring ejection fraction	0.00	90.24	Not Calculated
Hypertension	NA	85.37	66.67
Time from PTCA to surgery	100.00	66.67	NA
Left Main Stenosis	NA	85.37	34.13
Myocardial Infarction	NA	78.05	100.00
Myocardial Infarction (Yes/No)	NA	90.24	100.00
NYHA CHF Class	42.86	29.27	72.61
Operative Incidence	NA	90.24	9.45
PTCA on this Admission	NA	78.05	88.89
Peripheral Vascular Disease	NA	92.68	66.67
Gender	NA	100.00	NA
Status of patient at discharge	NA	100.00	Not Calculated
Ventricular Arrhythmia	0.00	92.68	100.00

## Sensitivity Analysis

In the second part of the analysis, the goal was to explore how the model fit, coefficients and final hospital ratings would be affected if audited data replaced the data submitted by the hospitals. Table 8 summarizes the effect of replacing audited information for CCMRP submissions that were audited, while leaving all other records unchanged. It is worth noting that the variables and coefficients contained in Table 8 do not match the final 1999 model. This is a result of several factors that led to both changes in the data used to produce the risk model and in the variables contained in the final model. The factors included: 1) the removal of CCS Class and NYHA Class variables from the final risk-adjustment model after determining that coding of these variables was completely unreliable; 2) a post-audit check of all hospitals' coding of discharge status using the OSHPD Patient Discharge Data; 3) a post-audit check of all hospital's accuracy of isolated CABG submissions (both under-submissions and over-submissions of possible non-isolated CABG procedures); and 4) additional data corrections

made by all hospitals including submission of new records, removal of old records, and corrections of status and other variables in response to coding concerns raised during the audit.

For the majority of variables, the change in the coefficient estimate and p-values is slight. However, the statistical significance and/or risk differential changed for the following variables: Creatinine, Dialysis, Diabetes, time between CABG and MI, and Acuity. Having Diabetes and a Myocardial Infarction 21-plus days ago become statistically significant in the edited model.

We also observed an interesting change in the *Acuity* variable. The risk differential of all higher acuity groups increased compared to the reference group. For instance, the estimated increase in risk of death for a salvage patient went from 11.71 times the risk of an elective patient in the model based on the original data to 22.46 in the model based on the edited data.

**Table 8: Impact on Fitted Model Characteristics when Replacing Audited Records with Information from Audit, 1999 Data**

		Model based on CCMRP Data			Model based on CCMRP Data and Audited Data Where Record was Audited		
		Estimate	p-value	OR*	Estimate	p-value	OR
Intercept		-7.74	0.00		-9.11	0.00	
Age (Years)		0.06	0.00	1.06	0.06	0.00	1.06
Gender	Female	Reference Group			Reference Group		
	Male	-0.50	0.00	0.61	-0.43	0.00	0.65
Race	White	Reference Group			Reference Group		
	Non-White	0.03	0.76	1.03	0.00	0.98	1.00
Creatinine (mg/dl)		0.18	0.00	1.20	0.01	0.15	1.01
Congestive Heart Failure		0.38	0.00	1.46	0.55	0.00	1.73
Hypertension		0.14	0.18	1.15	0.23	0.04	1.25
Dialysis		0.39	0.18	1.47	1.24	0.00	3.45
Diabetes		0.19	0.04	1.21	0.25	0.01	1.29
Peripheral Vascular Disease		0.24	0.02	1.28	0.21	0.06	1.23
Cerebrovascular Disease		0.23	0.04	1.26	0.26	0.02	1.30
Ventricular Arrhythmia		0.37	0.01	1.44	0.42	0.01	1.53
COPD		0.29	0.01	1.34	0.36	0.00	1.43
Operative Incidence	First Operation	Reference Group			Reference Group		
	Second Operation	0.84	0.00	2.31	0.86	0.00	2.36
	Third or Higher Operation	1.41	0.00	4.08	1.26	0.00	3.54
Myocardial Infarction	None	Reference Group			Reference Group		
	Yes, but when unknown	0.26	0.39	1.29	0.36	0.32	1.43
	21+ days ago	0.11	0.40	1.11	0.41	0.00	1.50
	7-20 days ago	0.66	0.00	1.93	0.65	0.00	1.91
	1-6 days ago	0.22	0.08	1.25	0.27	0.04	1.31
	Within 1 day	0.69	0.00	2.00	0.77	0.00	2.15
PTCA on this Admission		0.12	0.38	1.12	0.03	0.80	1.04

		Model based on CCMRP Data			Model based on CCMRP Data and Audited Data Where Record was Audited		
		Estimate	p-value	OR*	Estimate	p-value	OR
Angina	None	Reference Group			Reference Group		
	Stable	-0.15	0.40	0.86	-0.07	0.73	0.93
	Unstable	0.01	0.94	1.01	0.16	0.41	1.17
NYHA CHF Class	I	Reference Group			Reference Group		
	II	0.14	0.36	1.15	0.12	0.42	1.13
	III	-0.06	0.64	0.94	-0.17	0.22	0.84
	IV	0.26	0.04	1.30	0.17	0.18	1.19
CCS Angina Class	I	Reference Group			Reference Group		
	II	-0.29	0.16	0.75	-0.18	0.40	0.84
	III	-0.05	0.77	0.95	-0.13	0.47	0.88
	IV	-0.16	0.36	0.85	0.09	0.62	1.09
Acuity	Elective	Reference Group			Reference Group		
	Urgent	0.26	0.02	1.29	0.33	0.00	1.39
	Emergent	1.24	0.00	3.46	1.33	0.00	3.77
	Salvage	2.46	0.00	11.71	3.11	0.00	22.46
Ejection Fraction (%)		-0.02	0.00	0.98	0.00	0.00	1.00
Left Main Stenosis	50% or less	Reference Group			Reference Group		
	51% to 70%	0.15	0.25	1.16	0.01	0.96	1.01
	71% to 90%	0.17	0.26	1.18	0.39	0.00	1.47
	91% or more	0.27	0.16	1.31	0.34	0.08	1.41
Type of Coronary Disease	Single Vessel	Reference Group			Reference Group		
	Double Vessel	-0.21	0.30	0.81	-0.16	0.45	0.85
	Triple Vessel or More	-0.09	0.60	0.91	0.03	0.87	1.03
	Left Main Stenosis Only	-0.81	0.28	0.45	-0.72	0.34	0.49
Mitral Regurgitation	None	Reference Group			Reference Group		
	Trivial	-0.04	0.66	0.96	-0.12	0.23	0.89
	Mild	0.16	0.34	1.17	0.06	0.73	1.06
	Moderate	0.17	0.50	1.18	0.21	0.40	1.24
	Severe	0.19	0.73	1.21	0.29	0.60	1.34

Fit Statistics:

R <sup>2</sup>	0.188	0.202
c-statistic	0.818	0.833
Hosmer-Lemeshow $\chi^2$ (p-value)	9.303 (0.317)	23.068 (0.003)

The next question was whether the use of the audited data, and the subsequent changes in the model coefficients, resulted in changes to the hospital ratings. Table 9 shows the changes in ratings after replacing audited information for the information submitted to CCMRP. Note that

all records that were not audited are still included. The table is sorted in ascending order by the post-audit observed death percentage.

**Table 9: Rating Changes after Replacement with Audited Data, 1999 Data**

Hospital #	Pre-Audit				Post-Audit			
	Rating	Observed Death Rate	Expected Death Rate		Rating	Observed Death Rate	Expected Death Rate	
			Lower 95% CL	Upper 95% CL			Lower 95% CL	Upper 95% CL
5		0.00	0.00	7.23		0.00	0.00	6.35
27		1.03	0.43	8.28		1.04	0.00	6.88
30		1.11	0.74	4.42	O	1.11	1.88	5.54
77		1.16	0.58	2.75		1.16	0.83	3.05
50	O	1.00	2.09	5.61	O	1.25	1.52	4.68
17		1.38	1.02	2.24		1.38	1.17	2.39
69		2.36	0.30	5.87		1.60	1.35	6.32
49		1.93	1.78	4.39		1.77	1.65	4.21
64		1.99	1.60	4.36		1.81	1.68	4.37
59		2.55	0.64	5.10		2.04	0.40	4.56
60		2.07	0.94	4.65		2.07	1.04	4.82
52		2.15	1.90	4.76		2.15	1.92	4.63
24		2.55	1.13	4.39		2.28	1.59	4.98
32	•	4.58	0.00	3.47		2.29	0.16	4.09
35		2.46	1.19	4.75		2.48	0.99	4.43
76		2.48	0.25	4.91		2.50	0.16	4.69
39		2.95	0.18	3.55		2.53	0.52	3.80
1		3.13	0.00	5.02		3.13	0.37	6.01
81	•	3.18	0.07	2.85		3.18	0.82	3.66
43		3.36	0.48	3.60		3.38	0.75	4.00
16		3.47	0.63	4.91		3.47	1.05	5.30
22		3.58	1.50	5.03		3.58	1.83	5.36
33	•	5.56	0.00	4.57		3.74	0.00	5.40
79		3.78	3.19	9.49		3.78	1.47	6.89
21		3.93	3.12	9.38		3.93	1.88	7.41
31		4.52	0.92	5.86		3.98	3.47	8.20
54		4.02	0.52	4.13		4.02	1.65	5.42
20		4.05	1.45	7.24		4.07	1.14	6.70
12		4.10	0.59	4.16		4.12	0.88	4.62
57		4.31	0.14	6.56		4.35	0.00	5.65
44	•	5.38	0.00	5.27		4.65	1.47	6.85
7	•	5.59	1.68	4.55	•	5.59	2.07	4.93
34	•	5.97	0.00	5.59		5.97	0.40	6.93

Hospital #	Pre-Audit				Post-Audit			
	Rating	Observed Death Rate	Expected Death Rate		Rating	Observed Death Rate	Expected Death Rate	
			Lower 95% CL	Upper 95% CL			Lower 95% CL	Upper 95% CL
37	●	7.75	0.00	5.19	●	6.20	0.00	5.37
9	●	9.33	0.00	7.69		6.94	0.00	7.47
18		8.28	2.09	8.37	●	8.28	1.33	7.30

Notes:

- Hospital Performed Worse than Expected
- Hospital Performed Better than Expected

For hospital 30, the rating changed from “No Different than Expected” to “Better than Expected.” While the observed death percentage remained unchanged, the expected death range moved to the right. Closer inspection of the coding differences between the original CCMRP submission and the audited data indicated that this hospital down-coded several variables, among them *CCS Angina* and *NYHA CHF Class*.

Hospitals 32, 81, 33, 44, 34, and 9 were considered “Worse than Expected” performers pre-audit. Post-audit all of these hospitals were reclassified as “No Different than Expected.” For hospitals 32, 33, 44, and 9, the primary driver for this change was incorrect coding of a survivor as deceased. Correcting the discharge status information altered the observed death rate. For hospitals 81 and 34, the post-audit expected death range shifted to the right compared to the pre-audit range. As such the observed death rate no longer fell above the upper 95% confidence interval.

Finally, the classification status of hospital 18 changed from “No Different than Expected” pre-audit to “Worse than Expected” post-audit. This resulted from a shift left in the expected range of performance for this hospital compared to the pre-audit expected range. As such, the observed death rate using post-audit data fell above the upper 95% confidence interval.

## KEY FINDINGS FROM THE AUDIT OF HOSPITAL DATA

The goal of the audit was to verify the integrity of the hospitals’ data submissions to CCMRP, identify coding problems, identify implications of the missing value assignment, and confirm performance classifications.

- Concerning the integrity of the hospitals’ data submissions we found that most variables were coded acceptably. However, the audit found that *NYHA Class* and *CCS Class* were highly problematic variables. The inconsistent coding of *NYHA Class* for CHF functional status and *CCS Class* for angina functional status resulted in these variables as being unreliable for modeling purposes and these two variables were ultimately dropped from the final risk models for 1999 and All Quarters analyses.
- Other than problems with *NYHA Class* and *CCS Class*, there was no evidence of systematic coding problems for any other variables for hospitals that were classified as either “better” or “worse” than expected. However, we did observe that hospitals that were rated “worse” than expected tended to submit more missing values in their data and tended to down-code several variables. The audit provided important feedback to these facilities on their coding practices and the implications of their coding. CCMRP encouraged these hospitals to review all non-audited cases and to resubmit a corrected

file. The final performance rankings were based on the results of the audit and additional corrections submitted by hospitals.

- The submission of incomplete data by hospitals (i.e., missing information) may have serious implications for a hospital's performance ranking. This is especially true for important risk predictors such as acuity, where the audit found that the CCMRP policy of assigning the lowest risk category frequently did not match the severity of variable found in the medical record.<sup>12</sup> Missing information did not pose problems for variables that were "missing by design," such as the presence/absence of co-morbidities (e.g., Diabetes, Hypertension, Congestive Heart Failure). In these cases, the "missing value assignment" by CCMRP agreed with what was found in the medical record.
- The lack of agreement between hospital coding and the audited data for *Acuity* implied problems that are most likely associated with the subjective nature of coding this variable. Hospitals identified as having severe coding problems with this variable were asked to correct the coding prior to the final analysis. Given the high degree of disagreement in coding and the importance of this variable in the risk model, the definition of how to code acuity should be clarified for hospitals and this variable should be included in all future audits.
- The audit findings of high levels of disagreement for the outcome variable led to the decision to link the CCMRP data with the OSHPD PDD to further investigate potential discrepancies in discharge status across all hospitals participating in the program. This cross-check of all hospitals also focused on determining whether hospitals had submitted the correct number of isolated CABG cases (to evaluate under-submissions and submissions of non-isolated CABGs).
- The audit implied several changes in outlier status. These changes were largely due to incorrect coding of discharge status. Patients were coded as dead even though they were alive at the time of discharge. This finding evidenced the importance of coding the outcome variable correctly. As it is a rare event, even small changes can change the rating of a hospital.

---

<sup>12</sup> Note, CCMRP staff engaged in multiple contacts with each hospital in advance of the audit and post-audit to allow hospitals to fix missing data problems.